

EXHIBIT 3

JOSEPH SAVERI
LAW FIRM

601 CALIFORNIA STREET
SUITE 1000
SAN FRANCISCO CA 94108

TEL 415.500.6800
FAX 415.395.9940

Via Electronic Mail

September 20, 2023

Joseph C. Gratz
Morrison & Foerster LLP
425 Market Street
San Francisco
California 94105-2482
JGratz@mofo.com

**Re: Doe 1 v. GitHub, Inc, Lead Case No. 4:22-cv-06823-JST
OpenAI's Responses to Plaintiffs' First Sets of Requests for
Production and Interrogatories**

Dear Joe:

We are writing to confirm our understanding of OpenAI's positions on Plaintiffs' First Set of Requests for Production and First Set of Interrogatories following our September 5 & 7 meet-and-confer sessions. Please let us know if our understanding as described below is incorrect in any way by October 3, 2023.

Interrogatories

Verifications

The OpenAI entities have not provided verifications for their interrogatory responses and objections as required by Rule 33(b)(5). Plaintiffs request the OpenAI entities provide separate verifications signed by each party entity under oath immediately.

Specific Interrogatories

Interrogatories 1, 14:

Pursuant to Rule 33(d), OpenAI agrees to provide Bates numbers of documents produced in response to these interrogatories. You have not done so. Please do so without delay.

Joseph C. Gratz
September 20, 2023
Page 2

Plaintiffs understand that OpenAI is limiting its response to these interrogatories and responding only with regard to Codex. Plaintiffs disagree with OpenAI's limitation because agreements between Defendants and the identities of training licensors are discoverable, relevant, and admissible. We are happy to defer seeking Court intervention were you to promptly produce and identify documents for our review. Otherwise, we are at an impasse.

Interrogatory 2:

OpenAI refused to respond to this interrogatory on the basis of relevance. Plaintiffs disagree that this objection is appropriate. Plaintiffs believe the parties are at an impasse.

Interrogatory 3:

OpenAI refused to respond to this interrogatory on the basis of relevance. Plaintiffs disagree that this objection is appropriate. OpenAI indicated it may answer if the interrogatory were limited to individuals or entities that control one or more OpenAI entities. Plaintiffs accept that limitation. Please provide a prompt response. Plaintiffs make no representations about the adequacy of such a response and reserve all rights.

Interrogatory 4:

OpenAI refused to respond to this interrogatory on the basis of relevance. Plaintiffs disagree that this objection is appropriate. OpenAI indicated it would answer for the agreements identified in response to Interrogatory 1. As noted with respect to Interrogatory 1, you have not done so. Please do so without delay. Plaintiffs make no representations about the adequacy of such a response and reserve all rights. We note the open issues with respect to Interrogatory No. 1. See above. We reserve all rights with respect to those issues.

Interrogatory 5:

OpenAI provided three names in response to this interrogatory and represented that those individuals possess "cumulative knowledge" of everything about "engineering and development of Copilot and its

Joseph C. Gratz
September 20, 2023
Page 3

previous iterations, versions, or forms.” This is insufficient. Plaintiffs are entitled to a full and complete answer. There is no dispute this interrogatory calls for relevant information. Plaintiffs are entitled to the identity of persons with knowledge or possessing discoverable information regardless of whether more than one such person possesses the information. This interrogatory calls for the identification of such witnesses, not the information in their possession. Plaintiffs also note that the discovery process will determine whether information is cumulative or not. Plaintiffs request OpenAI provide a complete response including all responsive individuals. Unless OpenAI promptly provides a complete response, Plaintiffs believe the parties are at an impasse.

Interrogatories 6–7:

OpenAI refused to respond to these interrogatories as written on the bases of relevance and lack of knowledge. Plaintiffs disagree that these objections are appropriate. OpenAI indicated it may answer if the interrogatories were limited to identifying OpenAI-affiliated individuals. Plaintiffs accept that limitation. Please provide a prompt response. Plaintiffs make no representations about the adequacy of such a response and reserve all rights.

Interrogatory 8:

Plaintiffs accept OpenAI’s response insofar as it pertains to Codex. Plaintiffs make no representations about the adequacy of such a response and reserve all rights. Plaintiffs disagree with the exclusion of GPT-3 or GPT-4 information. Plaintiffs believe the parties are at an impasse regarding the relevance of information regarding GPT-3 and/or GPT-4, including with regard to OpenAI’s response to Interrogatories 13 and 14.

Interrogatory 9:

OpenAI refused to respond to this interrogatory on the basis that Copilot is not its product and thus its employees lack knowledge. That objection is not well taken. First, it does not appear that you have conducted an investigation to determine if OpenAI personnel possess responsive information. If they do, we are entitled to a response, even if your representations regarding ownership or control of Copilot are

Joseph C. Gratz
September 20, 2023
Page 4

accurate. Also, the interrogatory is not limited to OpenAI but more generally seeks the identification of knowledgeable persons, regardless of affiliation. Please provide a prompt response. If OpenAI does not, Plaintiffs believe the parties are at an impasse.

Interrogatory 10:

OpenAI refused to respond to this interrogatory on the basis of vagueness. Specifically, that it does not know what “Ethics Team” means. We are puzzled about this. Facts regarding this were specifically alleged in the complaint. Plaintiffs explained that “Ethics Team” in Interrogatory 10 refers to those involved with OpenAI’s “Safety & Alignment” projects, such as those mentioned or described in OpenAI blog posts tagged “Safety & Alignment,” which are listed here: <https://openai.com/blog?topics=safety-alignment>.

Interrogatory 11:

While OpenAI has provided limited responsive information, Plaintiffs are entitled to a full and complete answer. Plaintiffs clarified that this interrogatory seeks information about not just OpenAI software, but also third-party software used by OpenAI to engage in the activities set forth (“maintaining, supervising, managing, analyzing, programming, updating, troubleshooting, diagnosing, testing or modifying Copilot”), such as Slack, Teams, Azure, etc. OpenAI agreed to conduct further investigation and identify, at a minimum, responsive third-party products and will supplement its response after identifying them. Please provide an answer without delay. Plaintiffs make no representations about the adequacy of such a response and reserve all rights.

Interrogatory 12:

While Plaintiffs appreciate OpenAI responds that certain individuals referenced in Interrogatory 12 may be identified through OpenAI’s counsel, Plaintiffs are entitled to the information requested in the interrogatory. This information is of the type ordinarily requested and provided as a matter of course given its clear relevance, regardless of jurisdiction. *See, e.g.*, Form Interrogatories – General, available at <https://www.courts.ca.gov/documents/disc001.pdf>. Please supplement with that information at once.

Joseph C. Gratz
September 20, 2023
Page 5

Interrogatory 13:

OpenAI stated that responsive data is voluminous and asked that Plaintiffs consider an alternative to a full response, such as sampling. We think this is a productive suggestion. In addition, Plaintiffs stated that they also seek descriptions of the dataset (such as “all public code on GitHub available on X date”) in addition to a complete listing. Plaintiffs further note that while the corpus of documents used to train Codex may be large, a list of the documents should not be and would allow us to consider the volume issues you raise. Plaintiffs stand by this request and expect a complete response (limited to Codex and Copilot until the parties’ dispute regarding relevance of GPT-3 and GPT-4 is resolved).

Requests for Production of Documents

Preliminarily, as we stated on our meet and confers, Plaintiffs understand that each request falls broadly into one or more of three categories for collection and production: (1) “go-get” (i.e., producing documents within its possession, custody, or control); (2) traditional search term and custodian collection; or (3) a hybrid methodology. OpenAI cannot confirm which requests it will respond to by producing documents by which methodology. Further, while Plaintiffs understand that OpenAI has at least agreed to produce documents responsive to certain requests¹ (and has produced a smattering of documents), it has not informed Plaintiffs of the volume of the production nor provided a date for substantial completion.² For certain other requests, OpenAI has indicated that it will not produce documents³—for those requests, Plaintiffs believe we are at an impasse. Plaintiffs reserve all rights.

¹ E.g., Request Nos. 7–9, 13.

² Plaintiffs make no representations as to the adequacy of OpenAI’s production to date or its forthcoming production and reserve all rights to pursue appropriate follow-up after assessing OpenAI’s production.

³ Request Nos. 1, 3, 10, 12, 14–20, 22, 24, 26–33, 35–36, 39–43, 45–47, 51, 53, 56.

Joseph C. Gratz
September 20, 2023
Page 6

Global Issues

Training. For requests related to the training of Copilot, OpenAI has stated that it objects to responding because Copilot is a product not a model, and so was not trained. Without representing that Plaintiffs agree with OpenAI's characterization, Plaintiffs expect that if responsive documents related to training Copilot are maintained in OpenAI's possession, custody, or control, those documents would be produced.

Scope. For a number of requests,⁴ OpenAI objects to responding because certain products referenced in those requests (e.g., Azure, GPT-3) are not, in OpenAI's view, relevant or encompassed by this action. Plaintiffs disagree but Plaintiffs understand that OpenAI will not consider as responsive any documents related to these products. Plaintiffs believe the parties are at an impasse.

Specific Document Requests

For ease of reference, we refer below to abbreviated descriptions of each document request or cluster of document requests (e.g., "Request for Production 1: Copilot Individuals"). These shorthand references should not be construed as narrowing or otherwise altering the scope of the request as originally served by Plaintiffs.

Request for Production 1: Copilot Individuals

OpenAI has refused to produce documents responsive to this request because Copilot "was not trained" and was not created by OpenAI and thus its employees lack knowledge. This objection is not well taken. OpenAI must provide documents in its possession, custody, or control that identify any individuals it is aware of that are responsive to this request, regardless of whether they are directly affiliated with OpenAI and/or whether Copilot is an OpenAI product. If no such documents exist, then OpenAI should so respond.

⁴ E.g., Request Nos. 6, 23–24.

Joseph C. Gratz
September 20, 2023
Page 7

Requests for Production 2, 5–7: Organizational Charts

OpenAI has agreed to produce only organizational charts that include individuals involved with Codex—if any are located. As set forth above, Plaintiffs disagree that Codex is the only OpenAI product relevant to this case. Plaintiffs believe the parties are at an impasse regarding the relevance of information regarding other OpenAI products, such as GPT-3 and GPT-4. Plaintiffs also note that to the extent organizational charts for prior periods have been destroyed or not maintained, we would request the data or other native information from which the charts were generated or produced.

Requests for Production 3, 21–24: Agreements Between Defendants

OpenAI agreed to produce agreements between OpenAI and Microsoft regarding Codex or Copilot in response to certain of these requests but nothing else. Earlier iterations of agreements and communications, however, are also relevant and should be produced. The exclusion of such materials is not well taken. In addition, it is inappropriate for OpenAI to limit its production to only agreements with Microsoft regarding Codex or Copilot. Agreements with any entity regarding any products relevant to this case should be produced. Plaintiffs believe the parties are at an impasse regarding these requests.

Requests for Production 4, 25: Formation of OpenAI

OpenAI has agreed to produce only publicly available corporate formation documents. First, we do not know what you mean by that. Second, these requests seek information beyond public filings. Should OpenAI remain unwilling to provide more than just publicly available corporate formation documents, Plaintiffs believe the parties are at an impasse.

Request for Production 8: How Codex Works

OpenAI stated it will produce additional documents beyond just the Codex source code. Plaintiffs will review the documents produced and reserve their right to revisit this request if they determine the production is inadequate.

Joseph C. Gratz
September 20, 2023
Page 8

Requests for Production 9, 15–16, 37–39, 44, 46–49: Training Data, Including Selection, Licensing, and Use

OpenAI will only produce documents related to Codex, stated that responsive data is voluminous, and asked that Plaintiffs consider an alternative to a full response, such as sampling. Plaintiffs would like to take you up on this without delay. Please provide us with the sample. Plaintiffs agree to continue to meet and confer regarding production of the documents that constitute the training data itself, but Plaintiffs make no representations as to whether an alternative to a full response will be adequate. Plaintiffs reiterated that they also seek documents that describe and/or list the contents of the training dataset and they are entitled to that production. Plaintiffs stand by this request and expect a full and complete response.

Request for Production 10: Plaintiffs' Use of Codex

OpenAI requested that Plaintiffs inform OpenAI whether any of them used the Codex API and if so, which Plaintiffs used it. Plaintiffs agreed to confirm which Plaintiffs used the Codex API to allow OpenAI to perform a targeted search.

Requests for Production 11–14⁵: Integration of Codex into Copilot, including Training & Output Filtering

OpenAI agreed to produce the Codex source code in response to this request. Plaintiffs request that OpenAI also produce communications and other non-source code documents as well. OpenAI invited Plaintiffs to review the source code first and then request additional documents. While relevant, the production of source code is not fully responsive and does not satisfy the requests. Plaintiffs further reiterate their demand OpenAI produce all responsive documents to the request, not just the Codex source code, and that OpenAI produce any responsive information regarding Copilot, regardless of whether GitHub may possess the same information.

⁵ Plaintiffs clarified that “license test” in Requests 11–14 is a typo and should read “license text”.

Joseph C. Gratz
September 20, 2023
Page 9

Requests for Production 17–18: Codex Duplicate Output Filtering

OpenAI proposed providing information only for Plaintiffs' code. Plaintiffs do not agree that this satisfies the request. Nonetheless, Plaintiffs agree that this is an appropriate place to start. Please do so without delay. Plaintiffs reiterate their request that OpenAI ultimately produce all responsive documents. Plaintiffs expect that a full production will be made in a timely fashion.

Requests for Production 19–20: Copilot Duplicate Output Filtering

OpenAI stated that this information is only in GitHub's possession. Plaintiffs accept OpenAI's response but make no representations as to its adequacy at this juncture. Plaintiffs note GitHub's unwillingness and recalcitrance with respect to providing this information. Plaintiffs request confirmation OpenAI will take steps to ensure GitHub produces this information.

Requests for Production 26–27: OpenAI Ownership

OpenAI refused to produce documents in response to these requests on the bases of relevance and overbreadth. Plaintiffs agree to meet and confer further these requests. However, Plaintiffs demand that OpenAI produce documents sufficient to identify owners of each OpenAI defendant and percent of ownership for each owner by entity. Plaintiffs propose that Plaintiffs will review such documents and advise if additional documents are required.

Requests for Production 28–32: Documents Produced to Governmental Entities which Address or Refer to the Issues in this Case

OpenAI refused to produce documents. The objections are not well-taken. The parties are at an impasse.

Request for Production 33: Lobbying Efforts which Address or Refer to the Issues in this Case

OpenAI refused to produce documents. The objections are not well-taken. The parties are at an impasse.

Joseph C. Gratz
September 20, 2023
Page 10

Request for Production 34: Financial Documents

OpenAI agreed to produce documents containing information about its sales, revenue, and costs from 2021 to the present. Plaintiffs reserve the right to revisit this request should the production prove inadequate.

Request for Production 35: Agreements with Absent Class Members

OpenAI refused to produce responsive documents but stated it may agree to produce documents specific to Plaintiffs only. Please do so. Plaintiffs reserve the right to revisit this request should the production prove inadequate.

Requests for Production 40–43, 45: Azure documents Relevant to this Case

OpenAI refused to produce documents on the basis of relevance. Plaintiffs disagree and believe the parties are at an impasse.

Request for Production 50: Other Lawsuits

OpenAI refused to produce documents on the basis of relevance. Plaintiffs would agree to the production of complaints filed in federal or state court or complaints submitted to any federal or state law enforcement or regulatory agency.

Requests for Production 51–52: Advertising

OpenAI agrees that these Requests call for discoverable information. OpenAI refused to produce documents in response to this request on the basis that most responsive documents are public. This is not well-taken. Plaintiffs disagree with this basis for refusing to produce documents. OpenAI must produce responsive documents in its possession, custody, or control. Plaintiffs believe the parties are at an impasse.

Request for Production 53: Personnel Files of Select Employees

OpenAI refused to produce documents on the basis of relevance. Plaintiffs agreed to meet and confer further regarding appropriate narrowing.

Joseph C. Gratz
September 20, 2023
Page 11

Request for Production 54: Document Retention Policies

OpenAI agreed to produce responsive documents. Plaintiffs reserve the right to revisit this request should the production prove inadequate.

Requests for Production 55–56: Insurance

OpenAI will produce copies of responsive insurance policies. Plaintiffs take it from this Response that responsive documents exist. Plaintiffs disagree with OpenAI's proposed narrowing but will review OpenAI's production and reserve all rights to pursue additional responsive documents.

Request for Production 57: Communications Regarding Named Plaintiffs

OpenAI proposed to use search terms and custodians to narrow the scope of this request. Plaintiffs are amenable to this approach but reserve all rights to pursue further responsive documents should it prove inadequate.

Sincerely,


Joseph R. Saveri